TWO-SAMPLE ESTIMATION OF THE INTERGENERATIONAL ELASTICITY OF EXPECTED INCOME: THE IGETWOS COMMAND

Pablo A. Mitnik
(pmitnik@stanford.edu)

July, 2018

Abstract

Due to data constraints, the intergenerational income elasticity (IGE)—the workhorse measure of economic mobility—has very often been estimated with short-run income measures drawn from two independent samples and using the Two-Sample Two-Stage Least Squares (TSTSLS) estimator. The IGE conventionally estimated in the mobility literature, however, has been widely misinterpreted as pertaining to the conditional expectation of children's income when in fact it pertains to its conditional geometric mean. In line with recent work, this article focuses almost exclusively on the IGE of expected income. It (a) proposes that estimation of this IGE in the two-sample context be based on a recently advanced two-sample generalized method of moments (GMM) estimator of the exponential regression model, and (b) introduces the user-written program **igetwos**, which implements that estimator as well as a GMM version of the TSTSLS estimator. The new program allows to use Stata to estimate both the IGE of the expectation and the IGE of the geometric mean when the income information for parents and children is available in two independent samples.

Keywords: intergenerational income elasticity, economic mobility, elasticity of the expectation, exponential regression model, instrumental variables, generalized method of moments, two-sample estimation, two-sample two-stage least squares estimator.

1. Introduction

The intergenerational elasticity (IGE) is the workhorse measure of intergenerational economic mobility. It has been widely estimated over the last four decades, very often with the goal of providing a summary assessment of the level of income or earnings mobility within a country (for reviews, see Solon 1999:1778-1788; Corak 2006; Jäntti and Jenkins 2015:Secs. 10.5.2 and 10.5.3). Beyond this elementary goal, the IGE has been used, among other purposes, to conduct comparative analyses of economic mobility and persistence across countries, regions, demographic groups, cohorts, and time periods (e.g., Chadwick and Solon 2002; Hertz 2005, 2007; Aaronson and Mazumder 2008; Björklund and Jäntti 2000; Mayer and Lopoo 2008; Bloome and Western 2011); to examine the relationship between cross-sectional economic inequality and mobility across generations (e.g., Corak 2013; Bloome 2015); and to study the impact of social policies and political institutions on inequality of opportunity (e.g., Bratsberg et al. 2007; Landersø and Heckman 2016).

Now, despite the IGE's centrality in the intergenerational-mobility field, Mitnik and Grusky (2017) have recently shown that this elasticity has been widely misinterpreted. Indeed, the IGE has been construed as pertaining to the *expectation* of children's income conditional on their parents' income—as apparent, for instance, in its oft-invoked interpretation as a measure of regression to the (arithmetic) mean. However, the IGE estimated in the literature pertains to the conditional *geometric mean* of the children's income. As explained later, this not only makes all conventional interpretations of the IGE invalid but also has very deleterious methodological consequences.

Mitnik and Grusky (2017) have argued that both the conceptual and the methodological problems can be solved in a straightforward manner by simply replacing the IGE of the geometric mean (the de facto estimated IGE) by the IGE of the expectation (the IGE that mobility scholars thought they were estimating) as the workhorse intergenerational elasticity. They also called for effectuating such replacement, which requires identifying appropriate estimators and, when necessary, making them available in the statistical packages most frequently employed by mobility scholars.

As shown in recent work that has estimated the IGE of the expectation (Mitnik and Grusky 2017; Mitnik 2017a, 2017b and 2017c; Mitnik et al. 2015), if income information for both parents and children is included in the same sample, estimation of this IGE may be based on (a) the Poisson Pseudo Maximum Likelihood (PPML) estimator (Santos Silva and Tenreyro 2006), and (b) the additive-error version of the Generalized Method of Moments (GMM) instrumental variables (IV) estimator of the Poisson or exponential regression model (Mullahy 1997; Windmeijer and Santos Silva 1997), or GMM-IVP estimator for short. The PPML estimator may be used to estimate the IGE of expected income in any context in which the Ordinary Least Squares estimator has been or may be used to estimate the conventional IGE (Mitnik 2017a); similarly, the GMM-IVP estimator may be used to estimate the former IGE in any context in which the Two-Stage Least-Squares (or other linear IV estimators) have been or may be used to estimate the latter IGE (Mitnik 2017b). Both the PPML and the GMM-IVP estimators are available in Stata. The PPML estimator is available with several commands, including the command **poisson**. The GMM-IVP estimator is available with the command **ivpoisson**. Mitnik (2017d) provides a detailed tutorial for how to use these commands to estimate the IGE of expected income in Stata.

Unfortunately, relatively few countries count with samples in which the incomes of parents and their children are both measured in some period during adulthood. For this reason, the conventional IGE has very often been estimated not just with annual and other short-run proxy income measures but with short-run income measures drawn from two independent samples and using the Two-Sample Two-Stage Least Squares (TSTSLS) estimator (Jerrim et al. 2016). In order to make possible the estimation of the IGE of expected income when information is available in two independent samples, Mitnik (2007c) advanced a two-sample GMM estimator of the exponential regression model, to which I will refer as the GMM-E-TS estimator, as well as an associated generalized error-in-variables model. The GMM-E-TS estimator may be used to estimate the IGE of expected income in any context in which the TSTSLS estimator has been or may be used to estimate the conventional IGE.

In this article I introduce the user-written program **igetwos**, which implements the GMM-E-TS estimator in Stata and thus allows to easily estimate the IGE of the expectation with information from independent samples. This new program also implements a two-equation GMM version of the TSTSLS estimator. Although point estimates of the conventional IGE are straightforward to obtain with the TSTSLS estimator using available Stata commands (e.g., **regress**), conducting statistical inference requires accounting for the two-step nature of the estimation by using resampling approaches or complicated closed-form asymptotic variance estimators (see Inoue and Solon 2010) that are not available with any Stata command. This problem is solved by **igetwos**, which relies on standard asymptotic inferential procedures. In addition, while the TSTSLS estimator is not efficient when implemented with more than one instrument, its GMM counterpart is efficient regardless of the number of instruments.

The structure of the rest of the paper is as follows. I first explain why the conventional IGE pertains to the conditional geometric mean of children's income rather than to its conditional expectation, as well as Mitnik and Grusky's (2017) proposal to redefine the IGE used as the workhorse measure of economic mobility. Next, I introduce the GMM-E-TS estimator, and discuss its key properties when employed to estimate the IGE of expected income with long-run and with short-run proxy income variables, and with valid and invalid instruments. This discussion also includes some information regarding the "lifecycle biases" that may affect the estimation of IGEs. After that I present the command **igetwos**, and explain how to use it to estimate IGEs in the two-sample context. The last section offers brief concluding comments.

2. The IGE of what? Redefining the workhorse intergenerational elasticity

As I already indicated, the conventionally estimated IGE has been widely misinterpreted. While mobility scholars have interpreted it as the elasticity of the expectation of children's income or earnings conditional on parental income, that IGE pertains in fact to the conditional geometric mean. Closely following Mitnik and Grusky's (2017) analysis, the standard population regression function (PRF) posited in the literature, which assumes the elasticity is constant across levels of parental income, is:

$$E(\ln Y \,|x) = \beta_0 + \beta_1 \ln x, \qquad [1]$$

where Y is the child's long-run income or earnings, X is long-run parental income or father's earnings, β_1 is the IGE as specified in the literature, and I use expressions like "Z$|w$" as a shorthand for "$Z|W = w$." The parameter β_1 is not, in the general case, the elasticity of the conditional expectation of the child's income. This would hold as a general result only if $E(\ln Y|x) = \ln E(Y|x)$. But, due to Jensen's inequality, the latter is not the case. Instead, as $E(\ln Y \,|x) = \ln \exp E(\ln Y \,|x)$, and $GM(Y|x) = \exp E(\ln Y \,|x)$, Equation [1] is equivalent to

$$\ln GM(Y|x) = \beta_0 + \beta_1 \ln x, \qquad [2]$$

where GM denotes the geometric mean operator. Therefore, β_1 is the elasticity of the conditional geometric mean, i.e., the percentage differential in the geometric mean of children's long-run income with respect to a marginal percentage differential in parental long-run income.[1]

As the geometric mean is undefined whenever an income distribution includes zero in its support, the IGE is undefined as well when this is the case. Mitnik and Grusky (2017:Section IV) have argued that this has serious methodological consequences: It (a) makes it impossible to determine the extent to which parental economic advantages are transmitted through the labor market among women (as many women have zero earnings), and (b) greatly hinders research on the role that marriage plays in generating the observed levels of intergenerational persistence in family income (as many people remain single or have nonworking spouses, and therefore cannot be included in analyses examining the relationship between people's parental income and the income contributed by their spouses). As a result, the study of gender and marriage dynamics in intergenerational processes has been badly hampered. Equally important, Mitnik and Grusky (2017:Section III) have shown that, as a consequence of mobility scholars' expedient of dropping children with zero earnings from samples (to address what is perceived as the problem of the logarithm of zero being undefined), estimation of earnings IGEs with short-run proxy earnings measures is affected by substantial selection biases. This makes the current use of the IGE of men's individual earnings as an index of economic persistence and mobility in a country a rather problematic practice.[2]

To address these problems, Mitnik and Grusky (2017) have called for redefining the workhorse measure of economic mobility. This entails replacing the PRF of Equation [1] by a

[1] The parameter β_1 is (also) the IGE of the expectation only when the error term satisfies very special conditions (Santos Silva and Tenreyro 2006; Petersen 2017; Wooldridge 2002:17).

[2] Importantly, the methodological problems discussed in this paragraph can't be solved by replacing zeros by "small values" (Mitnik and Grusky, 2017:15-16).

PRF whose estimation delivers estimates of the IGE of the expectation in the general case. Under the assumption of constant elasticity, that PRF can be written as:

$$\ln E(Y|x) = \alpha_0 + \alpha_1 \ln x, \qquad [3]$$

where $Y \geq 0$, $X > 0$ and $\alpha_1 = \frac{d \ln E(Y|x)}{d \ln x}$ is the percentage differential in the expectation of children's long-run income with respect to a marginal percentage differential in parental long-run income. Crucially, (a) all interpretations incorrectly applied to the conventional IGE are correct or approximately correct under this formulation (see Mitnik and Grusky 2017:Section V.A), and (b) the IGE of the expectation is fully immune to the methodological problems affecting the IGE of the geometric mean and, in particular, is very well suited for studying the role of marriage in the intergenerational transmission of advantage (see Mitnik and Grusky 2017:Section V.B for details; and Mitnik et al. 2015:64-68 for an empirical application).

3. Estimation of the IGE of the expectation in the two-sample context

As I noted earlier, the conventional IGE has very often been estimated not just with annual and other short-run proxy income measures, but with short-run income measures drawn from two independent samples and using the TSTSLS estimator. The main concern with this strategy has been that the instruments typically available (e.g., parental education and occupation) are most likely positively correlated with both the logarithm of short-run parental income and the error term in the PRF of interest, making them endogenous.[3] When this is the case, estimates may still be useful if the sign of their asymptotic bias can be established. The standard interpretation in the literature has been that TSTSLS estimates are useful as upper-bound estimates. However, the available justification for this interpretation falls short, as it

[3] This has been a concern with the IV estimation of the conventional IGE more generally, regardless of whether it is estimated in the one-sample or in the two-sample context (Solon 1992; Mitnik 2017b; Jerrim et al. 2016).

counterfactually assumes that mobility scholars have long-run income variables in their samples (Jerrim et al. 2016), which is essentially never the case. A fully-correct analysis needs to take into account that estimation is based on short-run proxy measures, i.e., that there is measurement error. This applies equally to the conventional IGE and to the IGE of the expectation.

The two-sample GMM estimator of the exponential regression model advanced by Mitnik (2017c) is based on a two-sample two-step estimator. I present this estimator next. I first assume that the income variables are measured without error, and then introduce measurement error into the analysis. This is followed by a description of the approach for transforming the two-sample two-step estimator into a GMM estimator, a brief summary of what is known about the empirical performance of the estimator, and some comments regarding lifecycle biases.

3.1. The two-sample two-step estimator with a valid instrument

In the two-sample context, the "main sample" has the children's income information, the "auxiliary sample" has the parents' income information, and both samples have a common set of variables (e.g., parents' education, father's occupation) that may be used as instruments or predictors for the parents' income information. Operationally, the two-sample two-step estimator simply extends the approach used by the TSTSLS estimator of the linear regression model to the estimation of the exponential regression model. Thus, when there is only one instrumented variable (a) the first step estimates a linear projection of that variable on the instruments and any other right-side variables included in the second step, using information from the auxiliary sample, and (b) the second step estimates the exponential regression model of interest using the PPML estimator and the main sample, with the instrumented variable replaced by its predicted values (which are computed with the parameters estimated in the first step).

I specify in this subsection the assumptions under which this two-sample two-step estimator of the exponential regression model is consistent or approximately consistent when all income variables are measured without error.[4] Without any loss of generality, I assume in what follows that $E(Y) = E(\ln X) = 1$.[5]

Rewriting Equation [3] in additive-error form, the PRF of interest is

$$Y = \exp(\alpha_0 + \alpha_1 \ln X) + \Psi, \qquad [4]$$

where $E(\Psi|x) = 0$. Assuming for simplicity that there is only one quantitative instrument denoted by T (e.g., years of parental education), the fist-step equation is the following population linear projection:

$$\ln X = \gamma_0 + \gamma_1 T + R. \qquad [5]$$

Consider now the following assumptions, which apply to all *t* when relevant:

$A1.\ E(R|t) = 0$

$A2. \forall c > 0, E(\exp(cR)|t) = E(\exp(cR))$

$A2'. Var(R|\mathrm{t}) = \mathrm{Var}(R)$

$A3.\ \gamma_1 \neq 0$

$A4.\ E(\Psi|t) = E(\Psi).$

Assumptions A1 and A3 entail that the expectation of the logarithm of parental income conditional on the value of *T* is a linear function of that value, while assumptions A3 and A4 entail that *T* is a valid instrument. Assumption A2 is similar to the standard assumption made for the estimation of Poisson models with unobserved heterogeneity (see, e.g., Winkelmann 2008).

[4] Tersa et al. (2008) have shown that, in the nonlinear context, "predictor-substitution IV estimators" are not consistent in the general case.

[5] This involves no loss of generality because it can always be achieved by simply changing the monetary units used to measure income.

Assumption A2', which posits that the error in the first-step equation is homoscedastic, provides an alternative to A2. Although A2' is not strictly weaker than A2 (neither assumption entails the other), the fact that the dependent variable in the first-step equation is the logarithm of an income variable may make A2' more attractive than A2.

I start by showing that under assumptions A1, A2, A3 and A4 the two-sample estimator of α_1 is consistent. Substituting Equation [5] into Equation [4], and using A1 and A3, yields:

$$Y = \exp(\alpha_0 + \alpha_1 E(\ln X \,|T) + \alpha_1 R) + \Psi$$

$$E(Y|t) = \exp\big(\alpha_0 + \alpha_1 E(\ln X \,|t)\big)\, E(\exp(\alpha_1\, R)|t) + E(\Psi|t). \quad [6]$$

Using now A2, Equation [6] reduces to:

$$E(Y|t) = \exp\big(\alpha_0' + \alpha_1 E(\ln X \,|t)\big) + E(\Psi|t), \quad [7]$$

where $\alpha_0' = \alpha_0 + \ln E(\exp(\alpha_1 R)$; and it further reduces to

$$E(Y|t) = \exp\big(\alpha_0' + \alpha_1 E(\ln X \,|t)\big) \quad [8]$$

if assumption A4 also holds, that is, if the instrument is valid.[6] If the variable $E(\ln X \,|T)$ were available in the estimation sample, Equation [8] would be consistently estimated by the PPML estimator (e.g., Santos Silva and Tenreyro 2006). Under a standard identification condition for two-step M-estimators (e.g., Wooldridge 2002:354), the PPML estimator that replaces $E(\ln X \,|T)$ by consistent estimates obtained in the first step, is also consistent.

Going back to Equation [6], an alternative justification for this estimator as "approximately consistent" can be obtained by replacing A2 by A2'. Indeed, it can be shown (Mitnik 2017c:11) that this yields:

$$E(Y|t) \cong \exp\big(\alpha_0'' + \alpha_1 E(\ln X \,|t)\big) + E(\Psi|t), \quad [7']$$

[6] $E(\Psi|t) = 0$ follows from $E(\Psi|x) = 0$ (see Equation [4]) and assumption A4.

where $\alpha_0'' = \alpha_0 + \ln(1 + 0.5\ [\alpha_1]^2\ Var(R))$. Here I focus on the first justification, as both lead to pragmatically equivalent conclusions.

Resorting to Taylor-series expansions, Mitnik (2017a:16) has advanced an approximated closed form expression for α_1 in a PRF like [8]. For future reference, I note that it yields:

$$\alpha_1 \cong C_{\alpha_1} - \left[\left(C_{\alpha_1}\right)^2 - V_{\alpha_1}\right]^{\frac{1}{2}}, \qquad [9]$$

where

$$V_{\alpha_1} = 2\left[Var\big(E(\ln X\,|T)\big)\right]^{-1} \qquad [9a]$$

$$C_{\alpha_1} = [Cov(Y, E(\ln X\,|T)\,)]^{-1}.^{7} \qquad [9b]$$

3.2. The two-sample two-step estimator with an invalid instrument

Let's now assume that estimation is not based on the valid instrument T but on the invalid instrument $\boldsymbol{T}$, and that although A4 does not hold it is the case that:

$$A3'.\ \boldsymbol{\gamma_1} > 0$$

$$A4'.\ Cov(\Psi, \boldsymbol{T}) > 0.$$

(Throughout I use bold font to indicate that a parameter, expression or variable pertains to the analysis with the invalid instrument.) As $\boldsymbol{\gamma_1} > 0$ if and only if $\ln X$ and $\boldsymbol{T}$ are positively correlated, in the "long-run context" A3' and A4' are equivalent to the standard assumption that the (invalid) instruments typically available to mobility scholars are positively correlated with the logarithm of parental income and with the error term of the PRF of interest.

To determine the implications of A3' and A4', it is useful to rewrite the counterpart to Equation [7] as follows:

[7] Equations [9], [9a] and [9b] assume $E(Y) = 1$, which is true by hypothesis, and $E(E(\ln X\,|T)) = 1$. The latter follows from $E(\ln X) = 1$, which is true by hypothesis.

$$E(\boldsymbol{Y}|\boldsymbol{t}) = \exp\left(\boldsymbol{\alpha_0'} + \alpha_1 E(\ln X\,|\boldsymbol{t})\right) \qquad [10]$$

where $\boldsymbol{Y} \equiv Y - E(\Psi|\boldsymbol{T})$. Now, making use again of the approximated closed-form expression introduced above, we may write:

$$\alpha_1 \cong \boldsymbol{C}_{\alpha_1} - \left[\left(\boldsymbol{C}_{\alpha_1}\right)^2 - \boldsymbol{V}_{\alpha_1}\right]^{\frac{1}{2}}, \qquad [11]$$

where

$$\boldsymbol{V}_{\alpha_1} = 2\left[Var\left(E(\ln X\,|\boldsymbol{T})\right)\right]^{-1} \qquad [11a]$$

$$\boldsymbol{C}_{\alpha_1} = [Cov(\boldsymbol{Y}, E(\ln X\,|\boldsymbol{T})\,)]^{-1}$$

$$= \left[Cov\left(Y, E(\ln X\,|\boldsymbol{T})\right) - \boldsymbol{\gamma_1} Cov(\Psi, \boldsymbol{T})\right]^{-1}.[8] \qquad [11b]$$

Actual estimation, however, is not based on $\boldsymbol{Y}$ but on Y, which is equivalent to making $\boldsymbol{\gamma_1} Cov(\Psi, \boldsymbol{T}) = 0$. As assumptions A3' and A4' entail that $\boldsymbol{\gamma_1} Cov(\Psi, \boldsymbol{T}) > 0$, and $\frac{\partial \alpha_1}{\partial \boldsymbol{C}_{\alpha_1}} < 0$ (Mitnik 2017a:16), it follows that the probability limit of the two-sample two-step estimator of the IGE of the expectation with the invalid instruments typically available to mobility scholars is larger than the true parameter. This is the same conclusion that is obtained for the conventional IGE when the latter is estimated with the TSTSLS estimator, also under the assumption that the samples have information on long-run rather than short-run income.

3.3. Two-sample two-step estimation of the IGE of the expectation with short-run income variables

Let $Z \geq 0$ be the children's short-run income and $S > 0$ be the parents' short-run income. Without any loss of generality, I assume that $E(Z) = E(S) = 1$. And, just to simplify the

[8] Equations [11], [11a] and [11b] assume that $E(\boldsymbol{Y}) = 1$ and that $E\left(E(\ln X\,|\boldsymbol{T})\right) = 1$. This follows immediately from $E(Y) = 1$ and $E(\ln X) = 1$, which are true by hypothesis. In deriving [11b] I applied the law of total covariance. See Mitnik(2017c:12) for the step-by-step derivation.

exposition, I ignore here the possible existence of lifecycle biases, which the generalized error-in-variables model advanced by Mitnik (2017c) does consider (more on this in subsection 3.6).

The first-step equation is now the population linear projection

$$\ln S = \tilde{\gamma}_0 + \tilde{\gamma}_1 D + Q,$$

where D is a generic instrument, i.e., an instrument that may or may not be valid.

Given that lifecycle biases have been ruled out, the short-run income measures can be expressed as:

$$Z = Y + W \qquad [12]$$

$$S = X + P. \qquad [13]$$

From $E(Y) = E(X) = E(Z) = E(S) = 1$, it follows that $E(W) = E(P) = 0.$ Let's now make the following measurement-error assumptions:

$$M1.\, Cov(W, D) = 0$$

$$M2.\, E(P|d) = E(P).$$

It is easy to show that when these measurement-error assumptions hold (a) assumption A1 entails $E(Q|d) = 0$, and (b) each of assumptions A3 and A3' entails $\tilde{\gamma}_1 \neq 0$.

Therefore, under A1, M1 and M2, the second-step equation may be written as:

$$E(Z|d) = \exp(\tilde{\alpha}_0 + \tilde{\alpha}_1 E(\ln S\,|d)).$$

Then, using again the approximated closed-form expression employed before, we have:

$$\tilde{\alpha}_1 \cong C_{\tilde{\alpha}_1} - \left[\left(C_{\tilde{\alpha}_1}\right)^2 - V_{\tilde{\alpha}_1}\right]^{\frac{1}{2}}, \qquad [14]$$

where

$$V_{\tilde{\alpha}_1} = 2\left[Var\big(E(\ln S\,|D)\big)\right]^{-1}$$

$$= 2\,[Var(E(\ln X|D) + E(P|D))]^{-1}$$

$$= 2\,[Var(E(\ln X|D))]^{-1}. \qquad [14a]$$

$$C_{\tilde{\alpha}_1} = \left[Cov\big(Z, E(\ln S\,|D)\big)\right]^{-1}$$

$$= \left[Cov\big(Y, E(\ln X\,|D)\big) + Cov(Y, E(P|D)) + \tilde{\gamma}_1 Cov(W, D)\right]^{-1}$$

$$= [Cov(Y, E(\ln X|D))]^{-1}.^{9} \qquad [14b]$$

So let's assume that Equations A1, A2, A3 and A4 hold, that is, let's consider the case in which the instrument is valid. Comparing Equations [14], [14a] and [14b] with Equations [9], [9a] and [9b] makes clear that in this scenario the "short-run estimator" (the two-sample two-step estimator with short-run income variables) is a consistent estimator of the IGE of the expectation as long as the measurement-error assumptions $Cov(W, T) = 0$ and $E(P|d) = E(P)$ hold.

Let's consider next the case in which the instrument is invalid because A4 does not hold, but A3' and A4' do hold. Comparing now Equations [14], [14a] and [14b] with Equations [11], [11a] and [11b] shows that, under the same measurement-error assumptions, the short-run estimator is upward inconsistent with the invalid instruments typically available to mobility scholars.

3.4. Transforming the two-sample two-step estimator into a two-sample GMM estimator

In the one-sample context, and following the approach first advanced by Newey (1984), a two-step estimator—where the estimator in the second step is itself an M-estimator or a GMM estimator that depends on the first-step estimator—can be easily transformed into a two-equation GMM estimator (where the two equations are estimated simultaneously). To do so, the first-

[9] Equations [14], [14a] and [14b] assume that $E(Z) = 1$ and $E(E(\ln S\,|D)) = 1$. The former is true by hypothesis while the latter follows from $E(\ln S) = 1$, which is true by hypothesis.

order conditions for the two equations are “stacked,” so that the first-order conditions for the full GMM problem reproduce the first-order conditions of the estimators employed in each step.

There are two main advantages to using this approach. First, a two-equation GMM estimator only involves standard asymptotic inferential procedures, while a two-step estimator requires to account for the two-step nature of the estimation by using more complicated closed-form asymptotic variance estimators or resampling methods. Second, transforming a two-step estimator into a GMM estimator ensures efficient estimation (see Wooldridge 2002:425 and ff. for more details). As efficiency is achieved by weighting instruments in an optimal way, in finite samples the two-step and GMM estimators will produce identical estimates when there is only one instrument but will generally produce somewhat different estimates when there are multiple instruments.

With a small modification, the same approach can be used to transform the two-sample two-step predictor-substitution estimator of the exponential regression model introduced above into a two-sample two-equation GMM estimator. Let’s estipulate that the “auxiliary equation” is the equation from the first step, and the “main equation” is the equation from the second step. Then a two-sample GMM estimator of the IGE of the expectation, where the moment conditions are products of instruments and “modified residuals,” is obtained as follows: (a) replace the missing information in each of the samples—the logarithm of parents’ income in the main sample, children’s income in the auxiliary sample—by any value, e.g., zero, (b) stack the data from the two samples into one sample, adding an indicator variable to identify the observations from the auxiliary sample, (c) define the modified residuals entering the moment conditions associated to the auxiliary equation as the usual residuals multiplied by the indicator variable, (d) define the modified residuals entering the moment conditions associated to the main equation as the usual residuals multiplied by one minus the indicator variable, and (e) estimate the two-equation model by GMM in the usual way.

The key steps are (c) and (d). The modified residuals defined in those steps are equal to the usual residuals for the observations in the equation-dependent "relevant sample" but are always equal to zero—regardless of the value of the parameter vector—for the observations in the equation-dependent "irrelevant sample." Therefore, estimation can proceed as in the one-sample context. The resulting estimator is the GMM-E-TS estimator.

3.5. Empirical evidence

Mitnik (2017c) has offered empirical evidence strongly supporting the notion that estimates of the IGE of the expectation obtained with the GMM-E-TS estimator and the instruments typically available are upper-bound estimates. His results also show that instruments vary greatly with regards to the tightness of the upper bounds that they provide. For this reason, he suggested putting a good amount of effort into searching for "best invalid instruments" when using the GMM-E-TS estimator. This may involve looking for additional instruments beyond those typically employed by mobility researchers; using multiple instruments simultaneously, and possibly including interactions between them; and exploring the effects of alternative functional forms (e.g., entering an instrument in levels or in logarithms), as this has been shown to be very consequential in some contexts (Reiss 2016).

3.6. Lifecycle biases

Regardless of IGE concept and regardless of estimator, IGE estimates are susceptible to both left- and right-side lifecycle biases, i.e., biases that may result when the differences in short-run incomes between children or between parents do not capture well the differences in their long-run incomes (see, e.g., Mazumder 2005). Since the mid-2000s, the literature has relied on measurement-error models that allow for lifecycle biases and model their effects on estimates of the conventional IGE generated with the OLS and TSLS estimators (see Haider and Solon 2006; Mitnik 2017b; Nybom and Stuhler 2016). In three recent papers, Mitnik advanced functionally similar generalized error-in-variables models for the estimation of the IGE of the expectation with the PPML (Mitnik 2017a), GMM-IVP (Mitnik 2017b) and GMM-E-TS (Mitnik 2017c) estimators. All generalized error-in-variables models indicate that using short-run measures of economic status pertaining to specific ages should eliminate the bulk of lifecycle biases, while the empirical evidence available for both IGEs suggests that this should happen when parents' and children's income information is obtained close to age 40 (e.g.,Haider and Solon 2006; Nybom and Stuhler 2016; Mitnik 2017a, 2017b, and 2017c).

To simplify the exposition, in this section I have assumed that estimates are always free of lifecycle biases. Although this is not true, all conclusions drawn above still follow when the generalized error-in-variables model that does take lifecycle biases into account (see Mitnik 2017c) is substituted for the simpler measurement-model model I used, provided that both the children's and the parents' short-run income measures pertain to the "right points" of their lifecycles. So, as long as the children's and the parents' measures pertain to when they are close to 40 years old, lifecycle biases may be ignored, at least as a first approximation.

If this is not the case, however, those conclusions are unwarranted. This applies, in particular, to the conclusion that estimates of the IGE of the expectation based on the GMM-E-TS estimator (and short-run income measures) are upward biased with the instruments typically available. This may not be the case if children's income is measured when they are young enough and/or if parents' income is measured when they are old enough; in these cases, the downward lifecycle biases that result may more than compensate for the upward bias associated to the use of invalid instruments. Therefore, the simplified discussion of estimation biases in this paper can be used as reference as long as the short-run income measures pertain to the right ages of parents and children. If this is not the case, however, the more complicated analysis provided by Mitnik's (2017c) generalized error-in-variables model, and the associated empirical evidence, should be consulted to determine—to the extent that this is possible—the likely direction of any (net) estimation bias.

4. The igetwos command

4.1 Syntax

The syntax for the **igetwos** command is

igetwos *depvar* [*varlist1*] [*if*] [*in*] [*weight*], **instruments**(*varlist2*) **sampaux**(*varname1*)
depvaraux(*varname2*) [*options*]

where *varlist1* and *varlist2* may contain factor variables and fweights, iweights, and pweights are allowed.

4.2 Options

Required options

instruments(*varlist2*) specifies the instruments

sampaux(*varname1*) specifies the name of an indicator variable identifying the observations from the auxiliary sample

depvaraux(*varname2*) specifies the name of the log-parental-income variable, which is the dependent variable in the auxiliary equation

GMM weight matrix options

wmatrix(*wmtype*) specifies the weight matrix type; weight matrix type *wmtype* may be **robust** (the default), **cluster** *clustvar*, or **unadjusted**; these types of matrices are as defined in the Stata manual entry for the **gmm** command

winitial(*iwtype*) specifies the initial weight matrix; initial weight matrix *iwtype* may be **unadjusted**, **identity** (the default), or the name of a Stata matrix; these types of matrices are as defined in the Stata manual entry for the **gmm** command

SE/Robust options

vce(*vcetype*) specifies the type of standard error reported; *vcetype* may be **robust**, **cluster** *clustvar*, **bootstrap**, **jackknife**, or **unadjusted;** the default *vcetype* is based on the *wmtype* specified in the **wmatrix()** option; the types of standard errors, and the rules used to define the default *vcetype*, are as described in the Stata manual entry for the **gmm** command

Other options

nostandardize requires that, when estimating the elasticity of the expectation, the dependent variable in the main equation not be standardized by dividing it by its mean in the main sample (which normally helps the model converge); the default is to standardize this variable

noinitvalues requires that initial values not be provided for GMM estimation; the default is to provide initial values

show requires that the results of the regressions used to generate initial values be shown; the default is not to show them

technique(*optalg*) specifies the optimization technique to use; *optalg* may be **nr**, **bfgs**, **dfp**, **gn** (the default), or a combination of these algorithms, which are those used by the **gmm** command

othergmmoptions( ) can be used to specify any option allowed by the command **gmm** not listed above, as long as the option pertains to the interactive version of that command; for instance, specifying "**othergmm**(igmm igmmiterate(8) quickderivatives)" requires **igetwos** to use the iterative GMM estimator (doing up to 8 iterations), and to employ an alternative method of computing numerical derivatives for the variance-covariance matrix

geometricmean requires estimation of the elasticity of the conditional geometric mean; the default is to estimate the elasticity of the conditional expectation

altinitvalues requires that alternative initial values, pertaining to the two-step estimates of the elasticity of the conditional geometric mean, be provided for the GMM estimation of the elasticity of the conditional expectation; the default is to provide as initial values those pertaining to the two-step estimates of the elasticity of the conditional expectation

4.3. Description and remarks

The command **igetwos** estimates IGEs of children's income with respect to parental income in contexts in which (a) the measure of children's income (the dependent variable) is available in one sample (the main sample), the measure of parental income is available in a different sample (the auxiliary sample), and other parental variables that can be used as instruments (e.g., parental education, parental occupation) are available in both samples, and (b) the IGE is assumed to be constant across levels of parental income. By default, **igetwos** estimates the IGE of the conditional expectation of children's income, using the GMM-E-TS estimator introduced in Section 3. The command can also estimate the IGE of the conditional geometric mean of children's income, using a GMM version of the TSTSLS estimator, which requires specifying the option **geometricmean**; here the two-step procedure employed by the TSTSLS estimator is replaced by estimation of a two-equation model by GMM (relying on the approach described in Section 3.4 for the IGE of the expectation).

The program assumes that the names of variables used in estimation have been harmonized across samples, the auxiliary sample has been appended to the main sample, and there is an indicator variable, specified in **sampaux**(*varname*1), that is coded 0 for the observations in the main sample and 1 for the observations in the auxiliary sample. It also assumes that the dependent variable of the main equation, specified by *depvar*, is the children's income when estimating the IGE of the expectation and the logarithm of their income when estimating the IGE of the geometric mean. In addition, the program assumes that *depvar* is coded as 0 (rather than missing) in the auxiliary sample, and that the dependent variable in the auxiliary equation, specified by option **depvaraux**(*varname2*), is coded as 0 (rather than missing) in the main sample. If this is not the case, the program exits with an error message. At least one instrument needs to be specified with the option **instruments**(*varlist2*). Control variables (e.g., a polynomial on children's age), may be specified in *varilist1* if desired. If one of the samples includes weights (e.g., population weights) and the other doesn't, weights equal to 1 need to be generated in the latter sample. Otherwise, the program will exit with an error message if weighted estimation is requested.

The optimization algorithm employed sometimes makes a difference for whether the model converges or not, and for the time it takes to converge. If the model under estimation has difficulties converging, users should experiment with the various optimization algorithms available (and combinations thereof) by using the option **technique**( ). Combining 15 iterations of the modified Newton–Raphson algorithm followed by 5 of the Davidon–Fletcher–Powell algorithm—i.e., specifying **technique**(*nr* 15 *dfp* 5)—has proven particularly useful in some cases. However, estimation with this combination may be much slower than with the default Gauss–Newton method, so it should only be used if needed.

Specifying the option **othergmm**(*igmm*), which requests use of the iterative GMM estimator, may lead in some cases to estimates that are (non-negligibly) different. These estimates are likely to be more precise (Hall 2005: Sec. 2.4 and 3.6) than those obtained without that option. Specifying the option **othergmm**(*igmm igmmiterate*(#)), with # > 2 a small maximum number of iterations (e.g., # = 8), may be a good compromise if the iterative GMM estimator takes too long to converge.[10] The iterative GMM estimator with a maximum of # > 2 iterations can still be expected to be more efficient than the default (for which # = 2). The *wmtype* **unadjusted** and the *vcetype* **unadjusted** should in general not be used; they are allowed by **igetwos** mostly because they might serve a pedagogical purpose in some contexts.

The option **show** displays the estimates generated by the two-step estimator on which the relevant GMM estimator is based. By default, **igetwos** uses these estimates as initial values for the latter. In addition, when estimating the IGE of the expectation, the dependent variable is divided by its mean in the main sample—which affects the intercept of the linear predictor but not the elasticity—as this often facilitates convergence. Both behaviors can be changed by specifying the options **noinitvalues** and **nostandardize**, respectively. When the option **altinitvalues** is specified the two-step estimates pertaining to the IGE of the geometric mean are used as initial values for the GMM estimation of the IGE of expected income. This may be

[10] Moreover, there is no guarantee that the iterative GMM estimator will converge (e.g., Hall 2005:90).

useful in the very unlikely case that the two-step estimator of the IGE of the expectation has trouble converging.

In addition to what is stored by **gmm**, **igetwos** stores a few macros in e( ), which are listed in the command's help. The following standard postestimation commands are available: **lincom**, **nlcom**, **test** and **nltest**.

5. Estimating IGEs in the two-sample context: Examples

In this section I present examples of estimation of IGEs in the two-sample context, using the command **igetwos** and a subset of the data (from the Panel Study of Income Dynamics) employed by Mitnik (2017c). The "artificial two-sample" data I use represent U.S. men and women born between 1966 and 1974. The annual measure of parents' family income pertains to when the average age of the parents was close to 40 years old, while the short-run measure of children's income is an average of their family income when they were 35-38 years old.[11] I mostly focus on the estimation of the IGE of the expectation with the TS-GMM-E estimator, but at the end I also present an example of estimation of the IGE of the geometric mean with the GMM version of the TSTSLS estimator.

As required by **igetwos**, the names of the variables used in estimation are the same across samples, and the auxiliary sample has been appended to the main sample. The variables used in the examples are the following:

c_inc	Child's family income
c_ln_inc	Logarithm of child's family income
p_ln_inc	Logarithm of parental income
p_age	Average parental age
p_yeduc	Parents' total years of education
f_occ	Father's occupation
c_pweight	Sampling weights
cluster	Cluster variable

[11] Therefore, the income variables I use can be expected to generate very little lifecycle bias (see Section 3.6).

aux Indicator variable identifying the auxiliary sample

There is one more condition that the data must satisfy for use with **igetwos**: The children's income variables need to be coded as zero (rather than missing) in the auxiliary sample, and the parental income variable needs to be coded as zero (rather than missing) in the main sample. I start by reading the data and doing the needed replacements:

```
. use igetwos_data, clear

. qui replace c_inc = 0 if aux == 1

. qui replace c_ln_inc = 0 if aux == 1

. qui replace p_ln_inc = 0 if aux == 0
```

The command **igetwos** does not accept the survey prefix **svy** but does allow population weights, which need to be specified with the data at hand. It also allows to request a weight matrix that accounts for arbitrary correlation among observations within clusters, and that cluster-robust standard errors be computed, both of which are also needed here.[12] In all cases I specify the option **othergmm(nolog)** to save space, but it's better not to do so (to be able to examine the iteration log for potential convergence issues).

In the first example I use parents' years of education as instrument, producing an estimate of the IGE of expected income of 0.76. The command and the corresponding output are:

[12] This is needed because of the relationship between the observations in the main and the auxiliary samples (see Mitnik 2017c:22).

```
. igetwos c_inc [pw=c_pweight], instruments(p_yeduc) sampaux(aux) depvaraux(p_ln_inc) wmatrix (cluster cluster) othergmm(nolog)

Final GMM criterion Q(b) =  7.53e-31

note: model is exactly identified

GMM estimation

Number of parameters =   4
Number of moments    =   4
Initial weight matrix: Identity                   Number of obs   =      1,646
GMM weight matrix:     Cluster (cluster)

Two-sample estimation of the intergenerational elasticity of the expectation
GMM-E-TS estimator
                              (Std. Err. adjusted for 823 clusters in cluster)
------------------------------------------------------------------------------
             |               Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     p_yeduc |    .151169   .0118903    12.71   0.000     .1278644    .1744736
       _cons |   9.249505   .1560973    59.25   0.000      8.94356     9.55545
-------------+----------------------------------------------------------------
   /main_ige |   .7601429    .096159     7.91   0.000     .5716748     .948611
  /main_cons |  -8.522528   1.078968    -7.90   0.000    -10.63727    -6.40779
------------------------------------------------------------------------------
Instruments for equation aux: p_yeduc _cons
Instruments for equation main: p_yeduc _cons
The dependent variable was standardized by dividing it by its mean in the main sample
```

Mobility scholars often include polynomials on parents' or children's age as controls, or specify different intercepts by gender or birth cohort. In the next example the auxiliary equation and the linear predictor of the main equation include a quadratic polynomial on parents' age. As the model converges very slowly with the default algorithm, I specify the option **technique**(*nr* 15 *dfp* 5). I also conduct a Wald test of the null hypothesis that all parental-age coefficients are zero.

```
. igetwos c_inc c.p_age c.p_age#c.p_age [pw=c_pweight], instruments(p_yeduc) sampaux(aux) depvaraux(p_ln_inc) wmatrix (cluster
> cluster) tech(nr 15 dfp 5) othergmm(nolog)

Final GMM criterion Q(b) =  8.56e-31

note: model is exactly identified

GMM estimation

Number of parameters =   8
Number of moments    =   8
Initial weight matrix: Identity                   Number of obs   =      1,646
GMM weight matrix:     Cluster (cluster)

Two-sample estimation of the intergenerational elasticity of the expectation
GMM-E-TS estimator
                                 (Std. Err. adjusted for 823 clusters in cluster)
--------------------------------------------------------------------------------
                |               Robust
                |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
----------------+---------------------------------------------------------------
aux             |
          p_age |   .2168265   .0331215     6.55   0.000     .1519095    .2817436
                |
c.p_age#c.p_age |  -.0023949   .0003756    -6.38   0.000     -.003131   -.0016588
                |
        p_yeduc |   .1394899   .0109218    12.77   0.000     .1180835    .1608963
          _cons |    4.64172   .7400395     6.27   0.000     3.191269    6.092171
----------------+---------------------------------------------------------------
main_ige        |
          _cons |   .7979627   .1066729     7.48   0.000     .5888877    1.007038
----------------+---------------------------------------------------------------
main_other      |
          p_age |  -.0658816   .0593434    -1.11   0.267    -.1821926    .0504294
                |
c.p_age#c.p_age |    .000635   .0006781     0.94   0.349     -.000694    .0019641
                |
          _cons |  -7.347712   1.204494    -6.10   0.000    -9.708477   -4.986948
--------------------------------------------------------------------------------
Instruments for equation aux: p_yeduc p_age c.p_age#c.p_age _cons
Instruments for equation main: p_yeduc p_age c.p_age#c.p_age _cons
The dependent variable was standardized by dividing it by its mean in the main sample

. test [aux]p_age [aux]c.p_age#c.p_age [main_other]p_age [main_other]c.p_age#c.p_age

 ( 1)  [aux]p_age = 0
 ( 2)  [aux]c.p_age#c.p_age = 0
 ( 3)  [main_other]p_age = 0
 ( 4)  [main_other]c.p_age#c.p_age = 0

           chi2(  4) =   44.24
         Prob > chi2 =    0.0000
```

In all results tables produced by **igetwos**, the IGE of the expectation is identified by the term "ige." However, as we can see by comparing the output from the first two examples, the structure of the table is different when the equations include additional right-hand variables, beyond those that are required, and when they don't. In the last example the null hypothesis that all parental-age coefficients are zero is strongly rejected, while including parental-age controls increases the estimated IGE from 0.76 to 0.80.

In the above examples, the computation of cluster-robust standard errors is implied by the use of a weight matrix of the *wmtype* **cluster**. This is so because the default *vcetype* in **igetwos** is based on the *wmtype* specified in the **wmatrix**() option. Now, as the default *wmtype* is **robust,** the default vcetype is also **robust**. Therefore, had the data been the result of simple random sampling, estimation of the model in the first example could have been accomplished by the following command:

```
. igetwos c_inc, instruments(p_yeduc) sampaux(aux) depvaraux(p_ln_inc)
```

that is, without having to specify robust standard errors (which are mandatory with the GMM-E-TS estimator). This command only includes the minimum information required by **igetwos**: The children's income variable, one instrument, the indicator variable identifying the auxiliary sample, and the parents' income variable (in logarithms).

In the following example I estimate a model in which both parents' years of education and its square are used as instruments. In addition to the option **technique**(*nr* 15 *dfp* 5), I include in this case the options **igmm igmmiterate(8)** (as arguments for **othergmm**). This requires that the iterative GMM estimator rather than the default (two-step) GMM estimator be used, with a maximum of 8 iterations. When there is more than one instrument, use of the iterative GMM estimator is recommended unless it is too computationally expensive, as it is likely to improve precision in finite samples (e.g. Hall 2005:Sec. 2.4 and 3.6).[13] The results follow, showing a larger IGE estimate than in the previous two examples:

[13] When there is only one instrument the model is exactly identified, so requesting that the iterative GMM estimator be used would make no difference.

```
. igetwos c_inc [pw=c_pweight], instruments(c.p_yeduc c.p_yeduc#c.p_yeduc) sampaux(aux) depvaraux(p_ln_inc) wmatrix (cluster
>  cluster) othergmm(igmm igmmiterate(8) nolog) tech(nr 15 dfp 5)

note: iterative GMM parameter vector converged

Final GMM criterion Q(b) =  .0018108

GMM estimation

Number of parameters =   5
Number of moments    =   6
Initial weight matrix: Identity                   Number of obs   =      1,646
GMM weight matrix:     Cluster (cluster)

Two-sample estimation of the intergenerational elasticity of the expectation
GMM-E-TS estimator
                                     (Std. Err. adjusted for 823 clusters in cluster)
```

	Coef.	Robust Std. Err.	z	P>|z|	[95% Conf.	Interval]
p_yeduc	.2641736	.0691993	3.82	0.000	.1285455	.3998017
c.p_yeduc#c.p_yeduc	-.0045586	.0027657	-1.65	0.099	-.0099792	.0008621
_cons	8.575029	.4330422	19.80	0.000	7.726281	9.423776
/main_ige	.8392476	.1008401	8.32	0.000	.6416046	1.036891
/main_cons	-9.427891	1.129275	-8.35	0.000	-11.64123	-7.214553

```
Instruments for equation aux: p_yeduc c.p_yeduc#c.p_yeduc _cons
Instruments for equation main: p_yeduc c.p_yeduc#c.p_yeduc _cons
The dependent variable was standardized by dividing it by its mean in the main sample
```

In the last example of estimation of the IGE of the expectation I instrument the logarithm of parental income with father’s occupation:

```
. igetwos c_inc [pw=c_pweight], instruments(i.f_occ) sampaux(aux) depvaraux(p_ln_inc) wmatrix (cluster cluster) othergmm(igmm
>  igmmiterate(8) nolog)

note: maximum number of GMM iterations reached

Final GMM criterion Q(b) =  .0086305

GMM estimation

Number of parameters =  28
Number of moments    =  52
Initial weight matrix: Identity                   Number of obs   =      1,646
GMM weight matrix:     Cluster (cluster)

Two-sample estimation of the intergenerational elasticity of the expectation
GMM-E-TS estimator
                                                    (Std. Err. adjusted for 823 clusters in cluster)
----------------------------------------------------------------------------------------------------
                                         |               Robust
                                         |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----------------------------------------+----------------------------------------------------------
                                   f_occ |
                  Management Occupations |   .8471031   .2279306     3.72   0.000     .4003674    1.293839
         Business Operations Specialists |   .4077496   .2318577     1.76   0.079    -.0466831    .8621823
                   Financial Specialists |   .9509485   .2331198     4.08   0.000     .4940421    1.407855
 Computer and Mathematical Occupations   |   .9332384   .2413502     3.87   0.000     .4602007    1.406276
    Architecture and Engineering Occs.   |   .8612367   .2213648     3.89   0.000     .4273697    1.295104
       Life, Physical, and Soc. Sc. Occs.|   .9000847     .236265    3.81   0.000     .4370139    1.363155
         Community and Social Servs. Occs.|  .5672572   .2139008     2.65   0.008     .1480194    .9864951
                       Legal Occupations |   .9639085   .2043673     4.72   0.000     .5633559    1.364461
      Education, Train., and Libr. Occs. |   .8108536   .2153782     3.76   0.000     .3887202    1.232987
 Arts, Entert., Sports, and Media Occs.  |    .925338   .2440441     3.79   0.000     .4470203    1.403656
        Healthcare Pract. and Tech. Occs.|   1.600948   .3486855     4.59   0.000     .9175367    2.284359
          Protective Service Occupations |   .4813879     .230227    2.09   0.037     .0301512    .9326246
      Food Prep. and Serving Occupations |  -.5133608   .2324357    -2.21   0.027    -.9689264   -.0577952
          Build. and Grounds Maint. Occs.|   .4490451   .2706515     1.66   0.097    -.0814222    .9795123
        Personal Care and Service Occs.  |  -.7091729   .2726534    -2.60   0.009    -1.243564   -.1747821
                       Sales Occupations |   .8022835   .2194434     3.66   0.000     .3721824    1.232385
        Office and Admin. Support Occs.  |   .6616396   .2328086     2.84   0.004     .2053431    1.117936
     Farming, Fish., and Forestry Occs.  |  -.0956289   .2945765    -0.32   0.745    -.6729883    .4817305
                    Construction Trades  |   .3426118     .240438    1.42   0.154    -.1286379    .8138616
                      Extraction Workers |  -.1082022   .2137761    -0.51   0.613    -.5271958    .3107913
        Inst., Maint., and Repair Workers|    .666307   .2197162     3.03   0.002     .2356712    1.096943
                  Production Occupations |   .3118101   .2123924     1.47   0.142    -.1044712    .7280915
        Transp. and Material Moving Occs.|   .1863674   .2209405     0.84   0.399     -.246668    .6194027
           Military Specific Occupations |   .0890247   .2589855     0.34   0.731    -.4185774    .5966269
                        DK; NA; refused  |   .1579864   .2281778     0.69   0.489    -.2892339    .6052067
                                         |
                                   _cons |   10.62132   .2042851    51.99   0.000     10.22093    11.02172
-----------------------------------------+----------------------------------------------------------
                               /main_ige |   .7010749   .0851351     8.23   0.000     .5342131    .8679366
                              /main_cons |  -7.910494   .9483669    -8.34   0.000    -9.769259   -6.051729
----------------------------------------------------------------------------------------------------
Instruments for equation aux: 0b.f_occ 1.f_occ 2.f_occ 3.f_occ 4.f_occ 5.f_occ 6.f_occ 7.f_occ 8.f_occ 9.f_occ 10.f_occ
    11.f_occ 13.f_occ 14.f_occ 15.f_occ 16.f_occ 17.f_occ 18.f_occ 19.f_occ 20.f_occ 21.f_occ 22.f_occ 23.f_occ 24.f_occ
    25.f_occ 999.f_occ _cons
Instruments for equation main: 0b.f_occ 1.f_occ 2.f_occ 3.f_occ 4.f_occ 5.f_occ 6.f_occ 7.f_occ 8.f_occ 9.f_occ 10.f_occ
    11.f_occ 13.f_occ 14.f_occ 15.f_occ 16.f_occ 17.f_occ 18.f_occ 19.f_occ 20.f_occ 21.f_occ 22.f_occ 23.f_occ 24.f_occ
    25.f_occ 999.f_occ _cons
Warning: convergence not achieved
The dependent variable was standardized by dividing it by its mean in the main sample
```

As in the previous example with multiple instruments, here I requested that the iterative GMM estimator, with up to 8 iterations, be used (rather than the default GMM estimator, where the number of iterations is 2). Unlike in the previous example, however, in this case the GMM iterative estimator did not converge in 8 iterations, and this is noted in two places in the output: Indirectly at the top, where there is a note saying “maximum number of GMM iterations reached," and directly at the bottom, where there is a warning saying "convergence not achieved." This is fully unproblematic. Like the default estimator, any GMM iterative estimator based on # > 2 iterations is consistent; moreover, as already indicated, it is likely more efficient that the default estimator in finite samples.

As I am using data that are "artificially two-sample," and the original data include a measure of long-run parental income, the long-run IGE of the expectation can be estimated with the PPML estimator. This estimate is 0.60 (Mitnik 2017b: Table 3). Examining the results obtained in this section, we see that, as expected (a) all IGE estimates produced by the GMM-E-TS estimator are upper-bound estimates, and (b) there is substantial variation across instruments, i.e., while the estimate using father's occupation as instrument is 0.70, those obtained by using parents' years of education, alone and as second degree polynomial (0.76 and 0.84, respectively), provide looser upper bounds.

I finish the section with an example of estimation of the conventional IGE, in which I use parents' years of education as instrument. Estimation of the conventional IGE is achieved by substituting the logarithm of children's income for children's income as outcome variable, and specifying the option **geometricmean**. I also specify the option **show**, just to show that, with a single instrument, **igetwos** reproduces the point estimate obtained with the TSTSLS estimator.

```
. igetwos c_ln_inc [pw = c_pweight], instruments(p_yeduc) sampaux(aux) depvaraux(p_ln_inc) wmatrix (cluster cluster) geometri
> cmean show othergmm(nolog)

Generating initial values

First-step regression
(sum of wgt is   3.1493e+04)

Linear regression                               Number of obs     =        823
                                                F(1, 821)         =     161.24
                                                Prob > F          =     0.0000
                                                R-squared         =     0.1996
                                                Root MSE          =     .66842

------------------------------------------------------------------------------
             |               Robust
    p_ln_inc |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     p_yeduc |    .151169   .0119048    12.70   0.000     .1278015    .1745364
       _cons |   9.249505   .1562873    59.18   0.000     8.942735    9.556275
------------------------------------------------------------------------------

Second-step regression
(sum of wgt is   3.1455e+04)

Linear regression                               Number of obs     =        821
                                                F(1, 819)         =      75.33
                                                Prob > F          =     0.0000
                                                R-squared         =     0.1325
                                                Root MSE          =     .77855

------------------------------------------------------------------------------
             |               Robust
    c_ln_inc |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
p_ln_inc_hat |   .9117561   .1050486     8.68   0.000     .7055599    1.117952
       _cons |   1.060185   1.180002     0.90   0.369    -1.255999    3.376369
------------------------------------------------------------------------------

Joint estimation by GMM

Final GMM criterion Q(b) =  1.20e-31

note: model is exactly identified

GMM estimation

Number of parameters =   4
Number of moments    =   4
Initial weight matrix: Identity                   Number of obs   =      1,644
GMM weight matrix:     Cluster (cluster)

Two-sample estimation of the intergenerational elasticity of the geometric mean
GMM version of the TSTSLS estimator
                              (Std. Err. adjusted for 823 clusters in cluster)
------------------------------------------------------------------------------
             |               Robust
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     p_yeduc |    .151169   .0118903    12.71   0.000     .1278644    .1744736
       _cons |   9.249505   .1560973    59.25   0.000      8.94356     9.55545
-------------+----------------------------------------------------------------
/main_ige_gm |   .9117561   .1087229     8.39   0.000     .6986632    1.124849
  /main_cons |   1.060185   1.221688     0.87   0.386    -1.334279    3.454649
------------------------------------------------------------------------------
Instruments for equation aux: p_yeduc _cons
Instruments for equation main: p_yeduc _cons
```

The estimate of 0.91 is larger than the log-run estimate of the conventional IGE obtained with the underlying data and the OLS estimator, which puts that IGE at 0.70 (Mitnik 2017b: Table 3).

This is in agreement with the standard approach in the literature, which treats TSTSLS estimates as upper-bound estimates.

6. Concluding remarks

I have introduced the user-written program **igetwos**, which estimates IGEs in the two-sample context. Together, the already available commands **poisson** and **ivpoisson** and the new command **igetwos** offer a full suite of estimators of the IGE of the expectation. Using these estimators, it is possible to use Stata to easily estimate this IGE in all situations in which the IGE of the geometric mean has been estimated by mobility scholars. The first two commands implement the PPML and GMM-IVP estimators and may be used to estimate the IGE of the expectation with any dataset that may be used to estimate the conventional IGE with the OLS and TSLS estimators, respectively. Similarly, the new command introduced here may be used to estimate the IGE of the expectation with any pair of datasets that may be used to estimate the conventional IGE with the TSTSLS estimator. Given that the diffusion of statistical advances is closely tied to their embodiment in statistical software (Koenker and Hallock, 2001:153), and that Stata is extensively used by mobility scholars, I hope that **igetwos** will make a significant contribution to the replacement of the conventionally estimated IGE by the IGE of the expected income as the workhorse intergenerational elasticity.

7. References

Aaronson, David and Bhashkar Mazumder. 2008. "Intergenerational Economic Mobility in the US: 1940 to 2000." *Journal of Human Resources* 43(1): 139-172.

Björklund, Anders and Markus Jäntti. 2000. "Intergenerational Mobility of Socio-Economic Status in Comparative Perspective." *Nordic Journal of Political Economy* 26: 3-32.

Bloome, Deirdre. 2015. "Income Inequality and Intergenerational Income Mobility in the United States." *Social Forces*, 93(3): 1047-1080.

Bloome, Deirdre and Western. 2011. "Cohort Change and Racial Differences in Educational and Income Mobility." *Social Forces*, 90(2): 375-395.

Bratsberg, Bernt, Knut Røed, Oddbjørn Raaum, Robin Naylor, Markus Jäntti, Tor Eriksson and Eva Österbacka. 2007. "Nonlinearities in Intergenerational Earnings Mobility. Consequences for Cross-Country Comparisons." *The Economic Journal*, 117 (March) C72-C92.

Chadwick, Laura and Gary Solon. 2002. "Intergenerational Income Mobility among Daughters." *The American Economic Review* 92(1): 335-344.

Corak, Miles. 2006. "Do Poor Children Become Poor Adults? Lessons from a Cross Country Comparison of Generational Earnings Mobility." IZA Discussion Paper No. 1993, IZA.

Corak, Miles. 2013. "Income Inequality, Equality of Opportunity, and Intergenerational Mobility." *Journal of Economic Perspectives* 27 (3):79-102.

Haider, Steven and Gary Solon. 2006. "Life-Cycle Variation in the Association between Current and Lifetime Earnings." *American Economic Review* 96(4): 1308-1320.

Hall, Alastair. 2005. *Generalized Method of Moments*. Oxford: Oxford University Press.

Hertz, Tom. 2005. "Rags, Riches and Race: The Intergenerational Economic Mobility of Black and White Families in the United States." In *Unequal Chances. Family Background and Economic Success*, edited by Samuel Bowles, Herbert Gintis, and Melissa Osborne Groves. New York, Princeton and Oxford: Russell Sage and Princeton University Press.

Hertz, Tom. 2007. "Trends in the Intergenerational Elasticity of Family Income in the United States." *Industrial Relations* 46(1): 22-50.

Inoue, Atsushi and Gary Solon. 2010. "Two-Sample Instrumental Variables Estimators." *The Review of Economics and Statistics* 92(3):557-561.

Jäntti, Markus and Stephen Jenkins. 2015. "Income Mobility." In *Handbook of Income Distribution* (Vol. 2A), edited by Anthony Atkinson and François Bourguignon. Amsterdam: Elsevier.

Jerrim, John, Álvaro Choi, and Rosa Simancas. 2016. "Two-Sample Two-Stage Least Squares (TSTSLS) Estimates of Earnings Mobility: How Consistent Are They?" *Survey Research Methods* 10(2):85-102.

Koenker, Roger and Kevin Hallock. 2001. ''Quantile Regression.'' *Journal of Economic Perspectives* 15:143-156.

Landersø, Rasmus and James Heckman. 2016. "The Scandinavian Fantasy: The Sources of Intergenerational Mobility in Denmark and the U.S." NBER Working Paper 22465.

Mazumder, Bhashkar. 2005. "Fortunate Sons: New Estimates of Intergenerational Mobility in the United States Using Social Security Earnings Data." *The Review of Economics and Statistics* 87(2): 235-255.

Mayer, Susan E., and Leonard Lopoo. 2008. "Government Spending and Intergenerational Mobility." *Journal of Public Economics* 92: 139-58.

Mitnik, Pablo. 2017a. "Estimating the Intergenerational Elasticity of Expected Income with Short-Run Income Measures: A Generalized Error-in-Variables Model." Stanford Center on Poverty and Inequality Working Paper.

Mitnik, Pablo. 2017b. “Intergenerational Income Elasticities, Instrumental Variable Estimation, and Bracketing Strategies.” Stanford Center on Poverty and Inequality Working Paper.

Mitnik, Pablo. 2017c. “Two-Sample Estimation of the Intergenerational Elasticity of Expected Income.” Stanford Center on Poverty and Inequality Working Paper.

Mitnik, Pablo. 2017d. “Estimators of the Intergenerational Elasticity of Expected Income.” Stanford Center on Poverty and Inequality Working Paper.

Mitnik, Pablo, Victoria Bryant, Michael Weber and David Grusky. 2015. “New Estimates of Intergenerational Mobility Using Administrative Data.” SOI Working Paper, Statistics of Income Division, Internal Revenue Service.

Mitnik, Pablo and David Grusky. 2017. The Intergenerational Elasticity of What? A Call for Redefining the Workhorse Measure of Economic Mobility.” Stanford Center on Poverty and Inequality Working Paper.

Mullahy, John. 1997. “Instrumental-Variable Estimation of Count Data Models: Applications to Models of Cigarette Smoking Behavior.” *Review of Economics and Statistics* 79(4): 586-593.

Newey, Whitney. 1984. “A Method of Moments Interpretation of Sequential Estimators.” *Economics Letters* 14(2-3):201-206.

Nybom, Martin and Jan Stuhler. 2016. “Heterogeneous Income Profiles and Life-Cycle Bias in Intergenerational Mobility Estimation.” *The Journal of Human Resources* 15(1): 239-268.

Petersen, Trond. 2017. “Multiplicative Models for Continuous Dependent Variables: Estimation on Unlogged versus Logged Form.” *Sociological Methodology* 47:113-164.

Reiss, Peter. 2016, “Just How Sensitive are Instrumental Variable Estimates?”. *Foundations and Trends in Accounting* 10(2-4): 204–237.

Santos Silva, J. M. C. and Silvana Tenreyro. 2006. "The Log of Gravity." *The Review of Economics and Statistics* 88(4): 641-658.

Solon, Gary. 1992. “Intergenerational Income Mobility in the United States.” *American Economic Review* 82(3): 393-408.

Solon, Gary. 1999. “Intergenerational Mobility in the Labor Market.” *Handbook of Labor Economics,* Volume 3A, edited by Orley C. Ashenfelter and David Card. Amsterdam: Elsevier.

Terza, Joseph, Anirban Basu, Paul J. Rathouz. 2008. “Two-Stage Residual Inclusion Estimation: Addressing Endogeneity in Health Econometric Modeling.” *Journal of Health Economics* 27(3):531-543.

Windmeijer, F. A. G. and J. M. C. Santos Silva. 1997. “Endogeneity in Count Data Models: An Applications to Demand for Health Care.” *Journal of Applied Econometrics* 12: 281-294.

Winkelmann, Rainer. 2008. *Econometric Analysis of Count Data*. Berlin: Springer-Verlag.

Wooldridge, Jeffrey. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, Mass: The MIT Press.